Also, a type replacement facility was tested for updating primitive types such as such as integers that were not qualified with a size, and thus may change as the underlying architecture is upgraded. Finally, the designers leveraged a tool, called eclim, to bridge the gap between text editors and Integrated Development Environments (IDE), by running both of them simultaneously, and set up to communicate with each other. Through this mechanism, modern IDE features were made available through text editors, minimizing the learning curve for scientists already experienced with their conventions.

This work was done by Marc Abrams, Palabi Saboo, and Mike Sonsini of Harmonia Holdings Group, LLC for Goddard Space Flight Center. Further information is contained in a TSP (see page 1). GSC-16475-1

Automatic Data Filter Customization Using a Genetic Algorithm

This work predicts whether a retrieval algorithm will usefully determine CO_2 concentration from an input spectrum of GOSAT (Greenhouse Gases Observing Satellite). This was done to eliminate needless runtime on atmospheric soundings that would never yield useful results. A space of 50 dimensions was examined for predictive power on the final CO_2 results.

Retrieval algorithms are frequently expensive to run, and wasted effort defeats requirements and expends needless resources. This algorithm could be used to help predict and filter unneeded runs in any computationally expensive regime.

Traditional methods such as the Fischer discriminant analysis and decision trees can attempt to predict whether a sounding will be properly processed. However, this work sought to detect a subsection of the dimensional space that can be simply filtered out to eliminate unwanted runs. LDAs (linear discriminant analyses) and other systems examine the entire data and judge a "best fit," giving equal weight to complex and problematic regions as well as simple, clear-cut regions. In this implementation, a genetic space of "left" and "right" thresholds outside of which all data are rejected was defined. These left/right pairs are created for each of the 50 input dimensions. A genetic algorithm then runs through countless potential filter settings using a JPL computer cluster, optimizing the tossed-out data's yield (proper vs. improper run removal) and number of points tossed.

This solution is robust to an arbitrary decision boundary within the data and avoids the global optimization problem of whole-dataset fitting using LDA or decision trees. It filters out runs that would not have produced useful CO_2 values to save needless computation. This would be an algorithmic preprocessing improvement to any computationally expensive system.

This work was done by Lukas Mandrake of Caltech for NASA's Jet Propulsion Laboratory. For more information, contact iaoffice@jpl.nasa.gov.

This software is available for commercial licensing. Please contact Dan Broderick at Daniel.F.Broderick@jpl.nasa.gov. Refer to NPO-47788.

Tracker Toolkit

This software can track multiple moving objects within a video stream simultaneously, use visual features to aid in the tracking, and initiate tracks based on object detection in a subregion. A simple programmatic interface allows plugging into larger image chain modeling suites. It extracts unique visual features for aid in tracking and later analysis, and includes sub-functionality for extracting visual features about an object identified within an image frame.

Tracker Toolkit utilizes a feature extraction algorithm to tag each object with metadata features about its size, shape, color, and movement. Its functionality is independent of the scale of objects within a scene. The only assumption made on the tracked objects is that they move. There are no constraints on size within the scene, shape, or type of movement. The Tracker Toolkit is also capable of following an arbitrary number of objects in the same scene, identifying and propagating the track of each object from frame to frame. Target objects may be specified for tracking beforehand, or may be dynamically discovered within a tripwire region. Initialization of the Tracker Toolkit algorithm includes two steps: Initializing the data structures for tracked target objects, including targets preselected for tracking; and initializing the tripwire region. If no tripwire region is desired, this step is skipped. The tripwire region is an area within the frames that is always checked for new objects, and all new objects discovered within the region will be tracked until lost (by leaving the frame, stopping, or blending in to the background).

This work was done by Steven J. Lewis and David M. Palacios of Caltech for NASA's Jet Propulsion Laboratory. Further information is contained in a TSP (see page 1).

This software is available for commercial licensing. Please contact Dan Broderick at Daniel.F.Broderick@jpl.nasa.gov. Refer to NPO-48253.

Towards Efficient Scientific Data Management Using Cloud Storage

A software prototype allows users to backup and restore data to/from both public and private cloud storage such as Amazon's S3 and NASA's Nebula. Unlike other off-the-shelf tools, this software ensures user data security in the cloud (through encryption), and minimizes users' operating costs by using space- and bandwidth-efficient compression and incremental backup. Parallel data processing utilities have also been developed by using massively scalable cloud computing in conjunction with cloud storage.

One of the innovations in this software is using modified open source components to work with a private cloud like NASA Nebula. Another innovation is porting the complex backup-to-cloud software to embedded Linux, running on the home networking devices, in order to benefit more users.

This work was done by Qiming He of Open Research for Goddard Space Flight Center. For further information, contact the Goddard Innovative Partnerships Office at (301) 286-5810. GSC-16415-1